Markov Processes Relat. Fields **11**, 535–552 (2005)

Markov Processes and **Related Fields**

An Ising-Type Model for Spatio-Temporal Interactions

E. Järpe

Department of Mathematics, Physics and Electrical Engineering, P.O. Box 823, Halmstad University, SE 301 18 Halmstad, Sweden. E-mail: eric.jarpe@ide.hh.se
http://www.hh.se/staff/erja

Received May 5, 2004, revised February 10, 2005

Abstract. A model which possesses both spatial and time dependence is the Markov chain Markov field [7]. Here inference about the parameter for spatio-temporal interaction of a special case of a Markov chain Markov field model is considered. A statistic which is minimal sufficient for the spatio-temporal interaction parameter and its asymptotic distribution are derived. A condition for stationarity of the sufficient statistic process and the stationary distribution are given. Likelihood based inference methods such as estimation, hypothesis testing and monitoring are briefly examined.

KEYWORDS: asymptotic distribution, interaction, Markov chain Markov field, perfect simulation, stationarity, sufficient statistic

AMS SUBJECT CLASSIFICATION: Primary 37A60, 60J25, 82C22, Secondary 62M40, 92C55

1. Introduction

Spatio-temporal interaction models are relevant in a broad variety of application areas such as biometrics (epidemic propagation of infectious diseases), econometrics (influence of trading attitudes in a stock market), computer science (computer virus attacks in large computer networks), forestry (maintenance of large planted forests) and more. In the model of this paper the state of a site in a lattice is depending on the states of its nearest neighbours (in the Markovian sense) to an extent which is proportional to the global degree of clustering of the previous pattern. The interpretation will be further mentioned at the end of this section after having defined the model. The model was also considered for an ophthalmology study by Ibáñez and Simó [10].

In this section the model is formally presented and in Section 2 a statistic which is sufficient for the interaction parameter (the only parameter in the model) is derived. Inference regarding this parameter is made based upon the asymptotic distribution of the sufficient statistic which is treated in Section 3 and simulation perfectly according to the stationary distribution is briefly mentioned in Section 4. Inference matters such as estimation, hypothesis testing and monitoring are dealt with in Section 5, and finally in Section 6 the results are discussed. All proofs of results are deferred to the Appendix.

The spatio-temporal interaction model of this paper is a special case of the Markov chain Markov field [7]. Let S be a finite set consisting of n^2 positions, called sites, symbolically denoted by $i \in \{1, 2, \dots, n^2\}$, forming a finite square lattice in $\mathbf{Z}^2$.

The configuration space is a product space with a σ-algebra, $\mathcal{D}$, of all possible subsets of $D = \{-1, 1\}^{n^2}$. Let $X_{S,t} = \{X_{i,t} : i \in S\}$ be a random field on a probability space $(D, \mathcal{D}, \mathsf{P}_X)$. If $S' \subset S$, we denote $\{X_{i,t} : i \in S'\}$ by $X_{S',t}$. We consider the case where given the state of the neighbourhood at time t, $x_{\partial i,t}$, and the state of a statistic Q, q_{t-1}, (which is a function of the lattice state at time $t-1$, $x_{S,t-1}$), the random variable at site i and time t, $X_{i,t}$ is conditionally independent of all other sites at time t and all previous configurations $X_{S,s}$, $s \le t-2$. We assume that, for each $t \in \mathbf{Z}^+$, $X_{S,t}$ fulfils a positivity condition: $\mathsf{P}(X_{S,t} = x_{S,t}) > 0$, see [4]. For the sake of simplicity we denote $\{X_{i,u} : i \in S,\ 1 \le u \le t\}$ by $\{X\}_t$.

Definition 1.1. The sites i and j are called (*spatial, first order*) *neighbours*, denoted by $i \sim j$, if i and j are at unity Euclidean distance from each other. The *neighbourhood* of a site i is the set $\partial i = \{j \in S : i \sim j\}$.

To avoid edge problems, we map our square study region onto a torus. In practise, however, the size of edge effects vanishes as the lattice size $n \to \infty$. Due to the Hammersley – Clifford Theorem (see [4]), $\{X_{S,t} : t \in \mathbf{Z}^+\}$ is Gibbs distributed.

Definition 1.2. The *global conditional distribution* of the model of this paper is transition probability $p_X(x \mid x')$ in the Markov chain $X_{S,1}, X_{S,2}, \dots$. It is defined by

$$p_X(x \mid x') = p(X_{S,t} = x \mid X_{S,t-1} = x') = Z_{x'}^{-1} \exp\left(\varphi\, n^{-2} Q(x) Q(x')\right)$$

where $Z_{x'} = \sum_{y \in D} \exp(\varphi\, n^{-2} Q(y) Q(x'))$ is a normalising constant and $Q(x) = \sum_{i \sim j} x_i x_j$ is an energy function where summation with index $i \sim j$ means summing over all $i \in S$, $j \in \partial i : j < i$. φ is the spatio-temporal interaction parameter on some parameter space $\Phi \subseteq \mathbf{R}$.

Observe that $p_X(x \mid x') = p_X(x \mid q')$ whenever $q' = Q(x')$. The notation $Z_{q'}$ will be used alternatively for denoting the normalising constant $Z_{x'}$ where

$Q(x') = q'$. Since the lattice size is finite, the existence and uniqueness of a stationary distribution of X_t is guaranteed for all values of φ.

Theorem 1.1. *The process $\{X_{S,t} : t \in \mathbf{Z}^+\}$ is time-reversible and has stationary distribution $\pi_X(x) = c^{-1} Z_x$, where $c = \sum_{x\in D} Z_x$ is a normalising constant.*

Conditional on the states of the spatial neighbourhood of site i and the previous lattice pattern, the distribution of the random variable at site i is

$$\mathsf{P}(X_{i,t} = x_i \mid X_{\partial i,t} = x_{\partial i}, X_{S,t-1} = x') = Z_i^{-1} \exp\Big(\varphi\, n^{-2} Q(x') x_i \sum_{j\in\partial i} x_j\Big)$$

where $Z_i = 2\cosh(\varphi\, n^{-2} Q(x') \sum_{j\in\partial i} x_j)$ is the local normalising constant. Due to the positivity condition, this conditional probability takes its values on $(0, 1)$ for all values of the neighbourhood. Denoting the local distribution function, $\mathsf{P}(X_{i,t} = x_i \mid X_{\partial i,t} = x_{\partial i}, X_{S,t-1} = x')$, by $\rho_i(x_i)$, the model may be expressed as a logistic linear model

$$\log\Big(\frac{\rho_i(1)}{1-\rho_i(1)}\Big) = 2\varphi\, n^{-2} Q(x') \sum_{j\in\partial i} x_j.$$

The interpretation of this model may be that an individual of a society of some kind (maybe a voter in a public vote deciding to vote, say, democratic or republican, or an agent in a stock market deciding whether to buy or sell for instance) is influenced by the opinion of who he or she communicates to at the present time-point to the extent of the general degree of determination of the whole population at the previous time-step. When people/agents decide how to vote/trade, they are influenced by general mood of the whole society from previous time-steps, a mood which affects how much they themselves should depend on their own neighbours. The model might as well be relevant in forestry: trees in a planted forest might well be growing according to a square lattice in a *forestry disease model* — the binary states corresponding to a tree being "contaminated" or "not contaminated" where the tree disease spreads from tree to tree, present to a small extent when there is no epidemic, while an epidemic break-out is reflected by a change in the interaction parameter, or in *a forestry storm damage model* — the binary states corresponding to trees being "upright" or "fallen" since trees are more likely to fall if they are isolated (i.e. surrounded by fallen trees) rather than surrounded by upright trees and also by strong attraction in the previous time-step indicating strong wind which in turn implies that the states of the neighbouring trees are even more important.

2. Sufficient statistic

For many inference matters about the interaction parameter φ, we are satisfied with observations of a statistic which is sufficient for φ. Conditional on the previous state, Q_{t-1}, the current value, Q_t, is sufficient for φ (since the transition probability $p_Q(q \mid q')$ is in the exponential family). Unconditionally $\{Q\}_t$ is sufficient for φ.

It is easily checked that the state-space of Q_t is $D_Q = \{\pm 2n^2$ and $4\ell - 2n^2$, $2 \leq \ell \leq n^2 - 2\}$. Further, let us introduce the number $m_q = \#\{x \in D : Q(x) = q\}$ of distinct patterns, $x \in D$, that result in a certain state $q = Q(x)$.

Theorem 2.1. *$\{Q\}$ is a time-homogeneous Markov chain.*

As mentioned previously the lattice size is finite and thus the existence and uniqueness of a stationary distribution is guaranteed for all values of φ. However, for large lattices phase transition effects will become visible; for instance the variance of the statistic Q_t will become exceedingly large close to values of φ (namely $\pm 0.5 \log(\sqrt{2}+1)$) which correspond to points of phase transition in an infinite lattice.

Theorem 2.2. *$\{Q\}$ is time-reversible and has stationary distribution $\pi_Q(q) = c^{-1} m_q Z_q$ where $c = \sum_{q \in D_Q} m_q Z_q$ is a normalising constant (which is the same as in Theorem 1.1). When n is even, $e_\pi(Q_t) = 0$ regardless of φ.*

This result is striking in the sense that regardless of φ, $e_\pi(Q_t) = 0$ in contrast to the Ising model. Trivially $V_\pi(Q_t) = 2n^2$ if $\varphi = 0$.

3. Asymptotics

The expectation and variance of Q_t given $Q_{t-1} = q'$ for finite lattices are quite computationally intractable since for calculating the normalising constant one has to calculate the sum of 2^{n^2} terms. Instead a way to treat large lattice systems is by using approximations deriving from asymptotic behaviour.

Due to the relation between moments of Q_t conditional on $Q_{t-1} = q'$, the derivatives of the log normalising constant (see e.g. [14]) are

$$\begin{aligned}
\frac{d}{d\alpha} \log Z(\alpha)\big|_{\alpha=\varphi n^{-2} q'} &= e_\varphi(Q_t \mid Q_{t-1} = q'), \\
\frac{d^2}{d\alpha^2} \log Z(\alpha)\big|_{\alpha=\varphi n^{-2} q'} &= V_\varphi(Q_t \mid Q_{t-1} = q')
\end{aligned}$$

for finite lattice sizes n. When n is even S is a bigraph implying the following symmetry property.

Proposition 3.1. *When n is even $e_\varphi(Q_t \mid Q_{t-1} = q')$ is odd and $V_\varphi(Q_t \mid Q_{t-1} = q')$ is even in φ.*

Assuming $|\varphi| < 0.5 \log(\sqrt{2}+1)$ we have, due to the isotropy of the model, that there is a constant $\kappa \in (-2, 2)$ such that $\lim_{n\to\infty} q'/n^2 = \kappa$. Defining $\zeta(\alpha) = \lim_{n\to\infty}(1/n^2)\log\sum_{x\in D}\exp(\alpha\, Q(x))$ we have that [16]: as $n \to \infty$,

$$\frac{1}{n^2} e_\varphi(Q_t \mid Q_{t-1} = q') - \zeta'(\frac{\varphi q'}{n^2}) \to 0,$$

i.e.

$$\frac{1}{n^2} e_\varphi(Q_t \mid Q_{t-1} = q') \to \zeta'(\varphi\kappa)$$

and

$$\frac{1}{n^2} V_\varphi(Q_t \mid Q_{t-1} = q') - \zeta''(\frac{\varphi q'}{n^2}) \to 0,$$

i.e.

$$\frac{1}{n^2} V_\varphi(Q_t \mid Q_{t-1} = q') \to \zeta''(\varphi\kappa).$$

These limit functions are plotted in Figure 1 and some values are given in Tables 1 and 2.

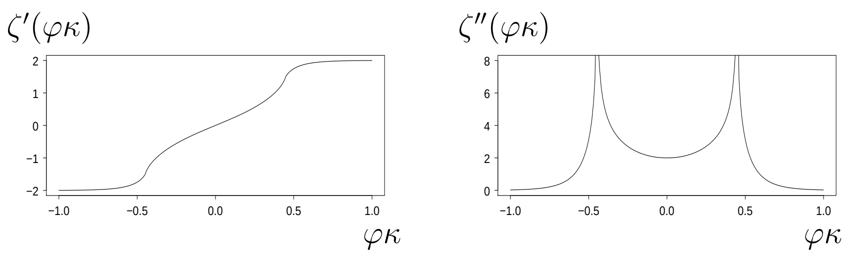

Figure 1. Asymptotic conditional moments of Q_t. Left picture: $\zeta'(\varphi\kappa) = \lim_{n\to\infty} n^{-2} e_\varphi(Q_t \mid Q_{t-1} = q')$ as a function of $\varphi\kappa$. Right picture: $\zeta''(\varphi\kappa) = \lim_{n\to\infty} n^{-2} V_\varphi(Q_t \mid Q_{t-1} = q')$ as a function of $\varphi\kappa$. In both pictures $\kappa = \lim_{n\to\infty} q'$.

Proposition 3.2. *For some* $\varepsilon \in (0, \varphi\kappa)$,

$$\lim_{n\to\infty} \frac{1}{n^2} e(Q_t \mid Q_{t-1} = q') = 2\varphi\kappa + \mathcal{O}(\varepsilon^3),$$
$$\lim_{n\to\infty} \frac{1}{n^2} V(Q_t \mid Q_{t-1} = q') = 2 + \mathcal{O}(\varepsilon^2).$$

Conditional asymptotic expected values of Q_t

$\varphi\kappa$	+0.00	+0.01	+0.02	+0.03	+0.04	+0.05	+0.06	+0.07	+0.08	+0.09
0.0	0.000	0.020	0.040	0.060	0.080	0.100	0.121	0.141	0.162	0.182
0.1	0.203	0.225	0.246	0.267	0.289	0.312	0.334	0.357	0.380	0.404
0.2	0.428	0.453	0.478	0.504	0.530	0.557	0.585	0.614	0.643	0.673
0.3	0.705	0.737	0.770	0.805	0.842	0.880	0.920	0.962	1.007	1.054

Table 1. Values of $\zeta'(\varphi\kappa) = \lim_{n\to\infty} n^{-2}e_\varphi(Q_t \mid Q_{t-1} = q')$ as a function of $\varphi\kappa$ where $\kappa = \lim_{n\to\infty} q'$.

Conditional variances of Q_t

$\varphi\kappa$	+0.00	+0.01	+0.02	+0.03	+0.04	+0.05	+0.06	+0.07	+0.08	+0.09
0.0	2.000	2.001	2.004	2.009	2.016	2.025	2.036	2.050	2.065	2.083
0.1	2.102	2.124	2.149	2.176	2.205	2.237	2.272	2.310	2.350	2.394
0.2	2.441	2.492	2.547	2.606	2.670	2.739	2.814	2.894	2.982	3.077
0.3	3.182	3.296	3.422	3.563	3.720	3.899	4.102	4.339	4.619	4.960

Table 2. Values of $\zeta''(\varphi\kappa) = \lim_{n\to\infty} n^{-2}V_\varphi(Q_t \mid Q_{t-1} = q')$ as a function of $\varphi\kappa$ where $\kappa = \lim_{n\to\infty} q'$.

Immediately from the central limit theorem of the Ising model (see [18]) we have that conditional on $Q_{t-1} = q'$,

$$\frac{Q_t - e_\varphi(Q_t \mid Q_{t-1} = q')}{\sqrt{V_\varphi(Q_t \mid Q_{t-1} = q')}} \longrightarrow N(0,1) \quad \text{as} \quad n \to \infty.$$

This means that approximately for large n, $Q_t \overset{\mathcal{D}}{=} N(n^2\zeta'(\varphi n^{-2}q'), n^2 \times \zeta''(\varphi n^{-2}q'))$ and thus, according to Proposition 3.2,

$$Q_t \overset{\mathcal{D}}{=} N(2\varphi q', 2n^2) \quad \text{conditional on} \quad Q_{t-1} = q'. \tag{3.1}$$

According to (3.1) we have for larger lattices the following result about stationary distribution of $\{Q\}$.

Theorem 3.1. *Approximately for large lattices*

$$Q_t \overset{\pi}{\in} N(0, n^2/(0.5 - 2\varphi^2)).$$

Further, (3.1) implies that approximately for large lattices the Markov chain $\{Q\}$ is asymptotically an $AR(1)$ process satisfying the recursive relationship $Q_{t+1} = 2\varphi Q_t + \varepsilon_t$ for all $t \in \mathbf{Z}^+$ where $\{\varepsilon_t\}$ is white noise with $\varepsilon_t \in N(0, 2n^2)$ and $Q_0 \in N(0, n^2/(0.5 - 2\varphi^2))$. Thus, for instance, $\{Q\}$ has covariance function $r_Q(\tau) = 2^{|\tau|+1}n^2\varphi^{|\tau|}/(1-4\varphi^2)$ and spectral density function $R_Q(f) = 2n^2/(1 + 4\varphi^2 - 4\varphi\cos(2\pi f))$, $-1/2 \leq f < 1/2$.

4. Perfect simulation

The technique to simulate exactly according to the stationary distribution is called *perfect simulation* (introduced by Propp and Wilson [19]). It is used in this paper with the modification for the anti-monotone case by Häggström and Nelander [9].

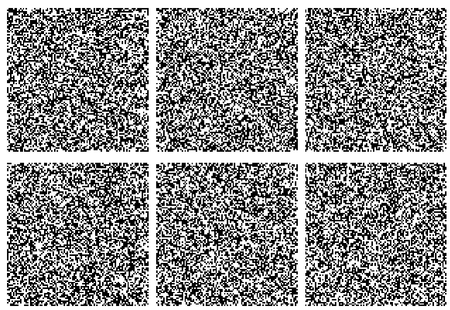

Figure 2. Three 100×100 square lattice patterns at times 1, 2, 3 simulated perfectly according to the stationary distribution of the spatio-temporal interaction model. In the top row $\{x\}_3^*$ with $\varphi=-0.3$, and in the bottom row $\{x\}_3^{**}$ with $\varphi=0.3$. Observe how hard it is to tell these patterns apart in respect of clustering by the mere eye. This stresses the importance of statistical methodology in order to distinguish clustering ($\varphi > 0$) or regularity ($\varphi < 0$).

The method of perfect simulation of the Ising model is applicable in the case with the spatio-temporal interaction model conditional on the initial lattice state. As well the initial q_0 state may be simulated using the approximate stationary distribution of $\{Q\}$ (see Theorem 3.1) and then the lattice state $x(0)$ may be chosen among $\{x \in D : Q(x) = q_0\}$ each with probability $m_{q_0}^{-1}$. Thus we may simulate samples according to the stationary distribution without bias introduced by approximating stationarity from sampling after a long run-in period. Apart from being nice for illustrations this way of getting perfect samples is important for the accuracy of results based on simulations. Two sequences, $\{x\}_{1000}^*$ and $\{x\}_{1000}^{**}$, were simulated with interaction parameter $\varphi = -0.3$ and $\varphi = 0.3$ respectively. In Figure 2 are the patterns $\{x\}_3^*$ and $\{x\}_3^{**}$ and in Figure 3 are the sequences $\{q\}_{1000}^*$ and $\{q\}_{1000}^{**}$ of the statistic $q_t = Q(x(t))$ plotted. The fact that the lattice is bipartite illustrated by the bottom image in Figure 3.

5. Inference

5.1. Estimation

Maximum likelihood estimation of the interaction parameter in the Ising model used to be regarded as an awkward task because of the tedious calculation of the normalising constant (e.g. [1] and [2]). Instead approximations (e.g. [15]) and other estimators, such as the pseudo-likelihood estimator and the generalised pseudo-likelihood estimator (e.g. [8] and [7]) have been suggested. The calculation of the approximations of maximum likelihood estimates is computationally heavy though, and the pseudo-likelihood estimator was shown biased towards underestimation in case of strong interaction [23]. However, for large lattices approximations using the asymptotic distribution and moments gives a good accuracy away from the points of phase transitions [11].

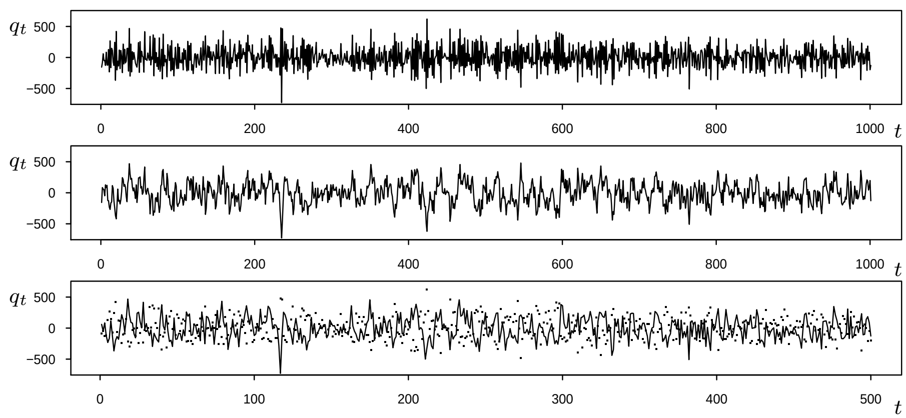

Figure 3. The two top plots show simulations $\{q\}^{*}_{1000}$ and $\{q\}^{**}_{1000}$ of $\{Q_t : t = 1, 2, \ldots, 1000\}$ based on simulations $\{x\}^{*}_{1000}$ and $\{x\}^{**}_{1000}$ of 100×100 square lattice sequences simulated perfectly according to the stationary distribution of the spatio-temporal interaction model. Values of time, t, are on the horizontal axis and corresponding values of the statistic, Q_t are on the vertical axis. In the top image $\varphi=-0.3$, and in the middle $\varphi=0.3$. Switching the sign of q_t deletes the effect introduced by a negative value of the interaction parameter. Thus a simulation of $\{Q_t\}$ for a value $\varphi < 0$ ($\varphi > 0$) may be obtained by simulating $\{Q_t\}$ for $\varphi > 0$ ($\varphi < 0$) and then switching the sign at even times $t = 2, 4, 6, \ldots$. The bottom plot shows the same simulation as above of $\{Q_t\}$ with $\varphi = 0.3$ except that it is plotted at odd times (indicated by a solid line) which thus is the same values as for the simulation with $\varphi = -0.3$. The dots indicate the values of $\{Q_t\}$ at even times.

Having made observations $x_0, x_1, \dots, x_t$ of $\{X\}$, the log likelihood function, under stationarity, is

$$l(\varphi) = \log\left(\pi_Q(Q(x_0)) \prod_{s=1}^{t} p_Q(Q(x_s)|Q(x_{s-1}))\right). \tag{5.1}$$

Using the observations $q_0, q_1, \dots, q_t$ thus yields the maximum likelihood estimator $\hat{\varphi}_{\mathrm{ML}}$ as the value of φ which maximises (5.1), i.e. the value of φ which solves the score equation

$$2\,\varphi\, e_\pi(Q_0^2) + \sum_{s=2}^{t} q_{s-1} e(Q_s \mid Q_{s-1} = q_{s-1}; \varphi) = \sum_{s=1}^{t} q_s q_{s-1}. \tag{5.2}$$

Example 5.1. Suppose that we are given the two simulated sequences $\{x\}^*_{1000}$ and $\{x\}^{**}_{1000}$ (of which $\{x\}^*_3$, $\{x\}^{**}_3$ and $\{q\}^*_{1000}$, $\{q\}^{**}_{1000}$ were shown in Figures 2 and 3 in Section 4) where 1: $\{x\}^*_{1000}$ was simulated with $\varphi = -0.3$ and 2: $\{x\}^{**}_{1000}$ with $\varphi = 0.3$. Using Proposition 3.2, approximate solutions to equation (5.2) are zeros of the score function

$$\frac{4\varphi n^2}{1 - 4\varphi^2} + 2\varphi \sum_{s=2}^{1000} q_{s-1}^2 - \sum_{s=1}^{1000} q_s q_{s-1} \tag{5.3}$$

with respect to φ. For the two cases the expression in (5.3) is plotted in Figure 4. From looking at values of this expression for different φ we have in four decimals accuracy:

1: $\hat{\varphi}^*_{\mathrm{ML}} = -0.2943$ in the case when $\varphi = -0.3$,

2: $\hat{\varphi}^{**}_{\mathrm{ML}} = 0.3025$ in the case when $\varphi = 0.3$.

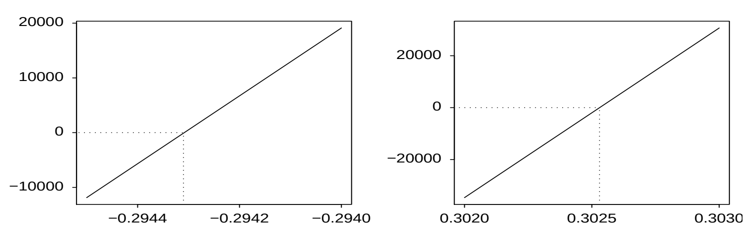

Figure 4. Plots of score functions, $\partial l/\partial\varphi$ (vertical axis), against φ (horizontal axis). The scores are based on samples $\{x\}^*_{1000}$ and $\{x\}^{**}_{1000}$ that were simulated with respectively $\varphi = -0.3$ (left picture), $\varphi = 0.3$ (right picture). The dotted lines indicate the estimates (i.e. zeros of the score) at -0.2943 and 0.3025 respectively.

5.2. Hypothesis testing

Since the sequence $\{Q\}_t$ is sufficient for φ, the log likelihood of $\{Q\}_t$ may be used as a test statistic for hypotheses regarding φ. Suppose that we want to test

$$\begin{cases} H_0: & \varphi = \varphi_0, \\ H_1: & \varphi \neq \varphi_0. \end{cases}$$

A test statistic based on the log likelihood function is

$$T(\varphi) = \frac{q_0^2(1-4\varphi^2)}{2n^2} + \sum_{s=1}^{t} \frac{(q_s - 2\varphi q_{s-1})^2}{2n^2} \tag{5.4}$$

for large n. This test statistic is approximately distributed $\chi^2(t)$ under H_0 and thus the null hypothesis should be rejected for large values of the statistic $T(\varphi_0)$. For the special case $\varphi_0 = 0$ the test statistic is reduced to $T(\varphi) = (2n^2)^{-1}\sum_{s=0}^{t} q_s^2$ where $T(\varphi) \in \chi^2(t_0)$ under $H_0 : \varphi = 0$ with approximations only due to using asymptotics for finite lattices.

Example 5.2. For the two sequences $\{q\}^*_{1000}$ and $\{q\}^{**}_{1000}$ plotted in Figure 3 in Section 4 we may test at level $\alpha = 0.01$ of significance

Test 1

$$\begin{cases} H_0 : \varphi = 0 \\ H_1 : \varphi \neq 0 \end{cases}$$ for the sequence $\{q\}^*_{1000}$ simulated with $\varphi = -0.3$,

Test 2

$$\begin{cases} H_0 : \varphi = 0.2 \\ H_1 : \varphi \neq 0.2 \end{cases}$$ for the sequence $\{q\}^{**}_{1000}$ simulated with $\varphi = 0.3$.

Here $n = 100$, $t = 1000$ and respectively for case 1: $\varphi_0 = 0$, 2: $\varphi_0 = 0.2$. Calculating the value of the test statistic $T(\varphi_0)$ according to (5.4) for these cases we get

1: $T^*(0) = 1547.2 > 1105.9 = \chi^2_{0.01}(1000)$ so $H_0 : \varphi = 0$ is rejected.

2: $T^{**}(0.2) = 1237.0 > 1105.9 = \chi^2_{0.01}(1000)$ so $H_0 : \varphi = 0.2$ is rejected.

5.3. Monitoring

There are many different approaches (monitoring, surveillance, change detection etc.) to the task of detecting a shift in a parameter of a distribution. Here we consider the case where the data accumulates in time and it is decided "on-line" whether or not a change has occurred. For further reading see e.g. [5, 6, 11, 12] or [22].

Consider the situation of lattice pattern observations, $X_{S,1}, X_{S,2}, \ldots$, being made consecutively with interaction parameter φ_0. At a random time-point, τ, the parameter changes to a new constant value, φ_1. In other words,

$$\mathsf{P}\left(X_{S,t} = x_t \mid X_{S,t-1} = x_{t-1}\right) = \begin{cases} p(x_t \mid x_{t-1}; \varphi_0) & \text{when } t \leq \tau, \\ p(x_t \mid x_{t-1}; \varphi_1) & \text{when } t > \tau. \end{cases}$$

The problem is to decide when this change occurs.

A method to this end is a stopping rule, e.g. $T = \inf\{t : a(\{X\}_t) > c\}$, where $a(\{X\}_t)$ is called alarm function and c threshold.

The Markov chain $\{Q\}_t$ is sufficient for φ so by using the approximations of Proposition 3.2 the approximate initial likelihood ratio under stationarity is $\mathrm{lr}(q_0^2 n^{-2}(\varphi_1^2 - \varphi_0^2))$ and the conditional likelihood ratio is $\mathrm{clr}(t) \approx \exp(q_{t-1}n^{-2} \times(q_t(\varphi_1 - \varphi_0) - q_{t-1}(\varphi_1^2 - \varphi_0^2)))$. According to this a method, corresponding to the frequently used *Cusum method* [17], is defined by the recursive relationship

$$a(\{q\}_t) = \begin{cases} (\log \mathrm{lr})^+ & \text{when} \quad t = 0, \\ (\log \mathrm{clr}(t) + a(\{q\}_{t-1}))^+ & \text{when} \quad t = 1, 2, 3, \ldots \end{cases}$$

A method, that corresponds to the *Shiryaev method* [21], is defined

$$a(\{q\}_t) = \begin{cases} \mathrm{lr} & \text{when} \quad t = 0, \\ \mathrm{clr}(t)(1 + a(\{q\}_{t-1})) & \text{when} \quad t = 1, 2, 3, \ldots \end{cases}$$

A method, corresponding to the *EWMA method* [20], is defined

$$a(\{q\}_t) = \begin{cases} \lambda \log \mathrm{lr} & \text{when} \quad t = 0, \\ \lambda \log \mathrm{clr}(t) + (1-\lambda) a(\{q\}_{t-1}) & \text{when } t = 1, 2, 3, \ldots, \end{cases}$$

where $\lambda \in (0, 1)$ is a weighting parameter. To get an idea of how well a monitoring method performs one may consider general quality measures such as expected time till a false alarm, $\mathrm{ARL}^0 = e(T \mid \tau = \infty)$, expected delay of a motivated alarm, $e(T - \tau \mid T \geq \tau = t)$, or probability of motivated alarm, $P(T = t \mid \tau = 1)$. Calibrating the thresholds of the Cusum, Shiryaev and EWMA (with $\lambda = 0.1$) methods such that $\mathrm{ARL}^0 = 100$, values of expected delay and values of probability of motivated alarm, when the interaction parameter changes from $\varphi_0 = 0$ to $\varphi_1 = 0.1$ and $\varphi_1 = 0.2$ respectively, are plotted in Figure 5. (These simulations were based on a sample size of 10 000 sequences of 100×100 patterns simulated by using perfect simulation.)

6. Discussion

Interaction systems in several dimensions sometimes possess instability properties that need to be included in the model for inference results to be correct. In this paper a spatio-temporal interaction model is examined. At large it is the proceedings of [11]. Applications could be e.g. in econometrics or forestry and it has recently been considered in an ophthalmology study by Ibáñez and Simó [10].

Means for perfect simulations of the spatio-temporal interaction model are indicated. It is shown that the sequence statistic which is sufficient for the spatio-temporal interaction parameter in the model, is a Markov chain. For large

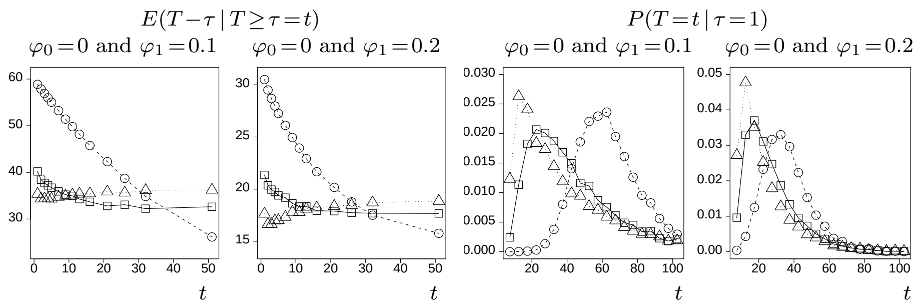

Figure 5. Expected delay and detection probabilities. In all four pictures are values of the conditional Cusum (squares and solid lines), Shiryaev (circles and dashed lines), and EWMA (triangles and dotted lines) methods. On the horizontal axes: time t. From left to right: the first two pictures show expected delay, $\mathsf{E}(T-\tau \mid T \geq \tau = t)$, and the last two show probability of motivated alarm, $\mathsf{P}(T=t \mid \tau=1)$. The first and the third pictures show change from $\varphi_0 = 0$ to $\varphi_1 = 0.1$, and the second and fourth pictures show change from $\varphi_0 = 0$ to $\varphi_1 = 0.2$.

lattices the stationary distribution and transition probabilities of this Markov chain are derived by approximations from asymptotics immediately accessible due to similarity to the Ising model. Methods for estimation, hypothesis testing and monitoring are indicated.

Further studies are needed for investigating the ergodicity and mixing properties. Adjustments of the model (such as e.g. adding inhomogeneity of spatio-temporal interaction, dividing the parameter space into regions interacting weakly or strongly) for being useful for potential applications, also remain for future studies.

Appendix.

Theorem A.1. *The process* $\{X_{S,t} : t \in \mathbf{Z}^+\}$ *is time-reversible and has stationary distribution* $\pi_X(x) = c^{-1}Z_x$, *where* $c = \sum_{x\in D} Z_x$ *is a normalising constant.*

Proof. Solving the equation $\pi_X(x)p_X(y \mid x) = \pi_X(y)p_X(x \mid y)$ with respect to π_X we have that $\pi_X(x)/Z_x = \pi_X(y)/Z_y$ since $\exp(\varphi n^{-2}Q(x)Q(y))$ is positive for all $x, y \in D$. This implies that $\pi_X(x) = c^{-1}Z_x$ where c is some constant with respect to x. Further $\sum_{y\in D} \pi_X(y) = c\sum_{y\in D} Z_y^{-1} = 1$ so let $c = \sum_{x\in D} Z_x$. Since $Z_x^{-1} > 0$ for all $x \in D$ and $c = \sum_{x\in D} Z_x^{-1}$, we have that $0 \leq \pi_X(x) = cZ_x^{-1} \leq 1$ for all $x \in D$. $\square$

Theorem A.2. $\{Q\}$ *is a time-homogeneous Markov chain.*

Proof. Let $\mathcal{X}_t = \{\{x\}_t \in D^t : Q(x_{S,1}) = q_1, \ldots, Q(x_{S,t}) = q_t\}$. Then

$$\begin{aligned}
p_Q\big(\{q\}_t \mid q_0\big) &= \sum_{x\in\mathcal{X}_t} p_X(x_{S,1} \mid q_0)\, p_X(x_{S,2} \mid x_{S,1}) \cdots p_X(x_{S,t} \mid x_{S,t-1}) \\
&= \sum_{x\in\mathcal{X}_t} Z_{q_0}^{-1} \exp\big(\varphi n^{-2} Q(x_{S,1}) q_0\big) \cdots Z_{q_{t-1}}^{-1} \exp\big(\varphi n^{-2} Q(x_{S,t}) q_{t-1}\big) \\
&= \frac{m_{q_1}}{Z_{q_0}} \exp\big(\varphi n^{-2} q_1 q_0\big) \cdots \frac{m_{q_t}}{Z_{q_{t-1}}} \exp\big(\varphi n^{-2} q_t q_{t-1}\big).
\end{aligned}$$

This implies that

$$\begin{aligned}
\frac{p_Q(\{q\}_t \mid q_0)}{p_Q(\{q\}_{t-1} \mid q_0)} &= \frac{m_{q_t}}{Z_{q_{t-1}}} \exp\big(\varphi n^{-2} q_t q_{t-1}\big) \\
&= \sum_{x:Q(x)=q_t} Z_{q_{t-1}}^{-1} \exp\big(\varphi n^{-2} Q(x) q_{t-1}\big) \\
&= p_Q(q_t \mid q_{t-1}).
\end{aligned}$$

Clearly, for any t and s, $p_Q(q_t \mid q_{t-1}) = p_Q(q_s \mid q_{s-1})$ whenever $q_t = q_s$ and $q_{t-1} = q_{s-1}$, which is to say that $\{Q\}$ is time-homogeneous. □

Lemma A.1. *If $\Psi : E \to F \subset \mathbf{R}$ is a function, E is finite, $\{Y\}$ is a Markov chain on E such that the transition probabilities satisfy the condition $p_Y(y \mid y') = p_Y(y \mid \Psi(y'))$ for all $y, y' \in E$, $\{Y\}$ has stationary distribution $\pi_Y(y)$, and $\{\Psi(Y)\}$ is a Markov chain, then $\{\Psi(Y)\}$ has stationary distribution $\pi_\Psi(\psi) = \sum_{y\in E:\Psi(y)=\psi} \pi_Y(y)$. If $\{Y\}$ is time-reversible, then so is $\{\Psi(Y)\}$.*

Proof. Since the sequence $\{Y\}$ is a Markov chain with stationary distribution $\pi_Y = [\pi_Y(e_1)\pi_Y(e_2)\ldots\pi_Y(e_N)]$, π_Y satisfies the unity eigenvalue equation $\pi_Y P_Y = \pi_Y$ where P_Y is the transition matrix of $\{Y\}$. Now, $p_Y(y \mid y') = p_Y(y \mid \Psi(y'))$ meaning that $\exists\, 0 \le N_1 < N_2 < \ldots < N_{k-1} \le N$ such that $p(e_\ell \mid e_{N_{j-1}+1}) = \ldots = p(e_\ell \mid e_{N_j})$ and $\Psi(e_{N_{j-1}+1}) = \ldots = \Psi(e_{N_j}) =: \Psi_j$ for $j = 1, \ldots k$, and $\ell = 1, \ldots, N$ where $N_0 := 0$ and $N_k := N$. Therefore the transition probabilities for $\Psi(Y)$ are (with $1 \le i, j \le k$)

$$p_\Psi(\psi_j \mid \psi_i) = \sum_{r=N_{j-1}+1}^{N_j} p_Y(e_r \mid e_s) \quad \text{whenever} \quad \psi_i = \Psi(e_s).$$

This defines the transition matrix, P_Ψ, of $\{\Psi(Y)\}$. Also $\Psi(e_\ell) \ne \Psi(e_m) \Rightarrow e_\ell \ne e_m$ since Ψ is a function. Thus, given that $\pi_Y P_Y = \pi_Y$, i.e.

$$\sum_{m=1}^{N} \pi_Y(e_m) p_Y(e_\ell \mid e_m) = \pi_Y(e_\ell) \quad \text{for each} \quad \ell, \tag{A.1}$$

we must show that π_Ψ is the eigenvector of P_Ψ. By using (A.1), we have

$$\begin{aligned}\sum_{i=1}^{k}\pi_\Psi(\psi_i)p_\Psi(\psi_j \mid \psi_i) &= \sum_{i=1}^{k}\Big(\sum_{r=N_{i-1}+1}^{N_i}\pi_Y(e_r)\Big)\Big(\sum_{s=N_{j-1}+1}^{N_j}p_Y(e_s \mid e_r)\Big)\\ &= \sum_{s=N_{j-1}+1}^{N_j}\Big(\sum_{i=1}^{k}\sum_{r=N_{i-1}+1}^{N_i}\pi_Y(e_r)\,p_Y(e_s \mid e_r)\Big)\\ &= \sum_{s=N_{j-1}+1}^{N_j}\pi_Y(e_s) = \pi_\Psi(\psi_j).\end{aligned}$$

To see that $\{\Psi(Y)\}$ is time reversible we need to check that

$$p_\Psi(\psi \mid \psi')\pi_\Psi(\psi') = p_\Psi(\psi' \mid \psi)\pi_\Psi(\psi) \quad \text{for all} \quad \psi, \psi' \in F.$$

Due to time reversibility of $\{Y\}$ we have that

$$\begin{aligned}p_\Psi(\psi \mid \psi')\pi_\Psi(\psi') &= \Big(\sum_{y\in E:\Psi(y)=\psi}p_Y(y \mid y')\Big)\Big(\sum_{y'\in E:\Psi(y')=\psi'}\pi_Y(y')\Big)\\ &= \sum_{y,y'\in E:\Psi(y)=\psi,\Psi(y')=\psi'}p_Y(y \mid y')\pi_Y(y')\\ &= \sum_{y,y'\in E:\Psi(y)=\psi,\Psi(y')=\psi'}p_Y(y' \mid y)\pi_Y(y)\\ &= \Big(\sum_{y'\in E:\Psi(y')=\psi'}p_Y(y' \mid y)\Big)\Big(\sum_{y\in E:\Psi(y)=\psi}\pi_Y(y)\Big)\\ &= p_\Psi(\psi' \mid \psi)\pi_\Psi(\psi).\end{aligned}$$

□

Lemma A.2. *Let the states of the statistic Q be denoted by $q_{(\ell)}$, $\ell = 1, \ldots, n^2-1$ and D_Q be the set of possible values of Q. If $n \geq 2$ is even, then $q_{(\ell)} = -q_{(n^2-\ell)}$ for all $\ell = 1, 2, \ldots, n^2-1$ and $m_q = m_{-q}$ for all $q \in D_Q$.*

Proof. Clearly $q_{(1)} = -2n^2 = -q_{(n^2-1)}$ and $2 \leq \ell \leq n^2-2$, $q_{(\ell)} = 4\ell - 2n^2$ so $q_{(n^2-\ell)} = 2n^2 - 4(n^2-\ell) = -2n^2 + 4\ell = -q_{(\ell)}$ for all integers $n \geq 2$.

For proving that $m_q = m_{-q}$ for even $n \geq 2$, we may consider the lattice S as a bigraph consisting of the sub-lattices S' and S'' defined so that $\forall i \in S'\ \exists!\, \partial i \subset S''$ and vice versa. This means that $S' = \{i : i \text{ odd (even) at odd (even) rows}\}$ and $S'' = \{i : i \text{ even (odd) at odd (even) rows}\}$. Let $q_{(\ell)} \in D_Q$ be given and let

$x_{(1)}, \ldots, x_{(r)}$ be an enumeration of $D_\ell = \{x \in D : Q(x) = q_{(\ell)}\}$ (implying that $r = m_{q_{(\ell)}}$). Let $\mathcal{T} : D \to D$ be the map defined by $\mathcal{T}(x)_{(k)} = (x_{(k)})_{S'} - (x_{(k)})_{S''}$, where $1 \leq k \leq r$, and construct the sequence $\mathcal{T}(x)_{(1)}, \ldots, \mathcal{T}(x)_{(r)}$. Then

$$\begin{aligned} Q(\mathcal{T}(x)_{(k)}) &= \sum_{i \in S'} \sum_{j \in \partial i : j > i} (x_{(k)})_i (-x_{(k)})_j + \sum_{i \in S''} \sum_{j \in \partial i : j > i} (-x_{(k)})_i (x_{(k)})_j \\ &= -\sum_{i \in S} \sum_{j \in \partial i : j > i} (x_{(k)})_i (x_{(k)})_j \\ &= -Q(x_{(k)}) \end{aligned}$$

There cannot be a pattern, $\mathcal{T}(x)_{(r+1)}$ such that $Q(\mathcal{T}(x)_{(r+1)}) = -q_{(\ell)}$, because then $(\mathcal{T}(x)_{(r+1)})_{S'} - (\mathcal{T}(x)_{(r+1)})_{S''}$ should have been in D_ℓ, but $x_{(1)}, \ldots, x_{(r)}$ is an enumeration of D_ℓ. Hence $\mathcal{T}(x)_{(1)}, \ldots, \mathcal{T}(x)_{(r)}$ is an enumeration of $\{x \in D\colon Q(x) = -q_{(\ell)}\}$ and thus

$$m_{q_{(\ell)}} = \# D_\ell = \#\{x \in D : Q(x) = -q_{(\ell)}\} = m_{-q_{(\ell)}}.$$

□

Theorem A.3. $\{Q\}$ *is time-reversible and has stationary distribution* $\pi_Q(q) = c^{-1} m_q Z_q$ *where* $c = \sum_{q \in D_Q} m_q Z_q$ *is a normalising constant* (*the same as in* Theorem 1.1). *When* n *is even,* $e_\pi(Q_t) = 0$ *regardless of* φ.

Proof. Since the map $Q\colon D \to D_Q$ is a function, it is immediate from Lemma A.1 that $\pi_Q(q) = \sum_{x : Q(x) = q} \pi_X(x)$ is the stationary distribution of $\{Q\}$ where $\pi_X(x)$ is the stationary distribution of $\{X\}$ in Theorem 1.1. To prove that even n implies $e_\pi(Q_t) = 0$, write

$$Z_q = m_0 + 2 \sum_{\ell=1}^{n^2/2-1} m_{q_{(\ell)}} \cosh(\varphi n^{-2} q q_{(\ell)}).$$

Since both cosh and m_q are even in q, by Lemma A.2, Z_q is as well and thus

$$\begin{aligned} e_\pi(Q_t) &= \sum_{q \in D_Q} q \pi(q) = \sum_{\ell=1}^{n^2-1} q_{(\ell)} c^{-1} m_{q_{(\ell)}} Z_{q_{(\ell)}} \\ &= c^{-1} \sum_{\ell=1}^{n^2/2-1} q_{(\ell)} m_{q_{(\ell)}} Z_{q_{(\ell)}} + c^{-1} \sum_{\ell=1}^{n^2/2-1} q_{(n^2-\ell)} m_{q_{(n^2-\ell)}} Z_{q_{(n^2-\ell)}} \\ &= 0, \end{aligned}$$

since $q_{(n^2/2)} = 0$, $q_{(\ell)} = -q_{(n^2-\ell)}$, $m_q = m_{-q}$ (when n is even) and $Z_q = Z_{-q}$ for $1 \leq \ell \leq n^2 - 1$ and for all $q \in D_Q$. □

Proposition A.1. *When n is even $e_\varphi(Q_t \mid Q_{t-1} = q')$ is odd and $V_\varphi(Q_t \mid Q_{t-1} = q')$ is even in φ.*

Proof. We need show that $e_\varphi(Q_t \mid Q_{t-1} = q') = -e_{-\varphi}(Q_t \mid Q_{t-1} = q')$. The map $\mathcal{T}$ in the proof of Lemma A.2 satisfies $Q(\mathcal{T}(x)) = -Q(x)$. Therefore

$$\begin{aligned}
e_\varphi(Q_t \mid Q_{t-1} = q') &= \sum_{x \in D} Q(x) Z_{q'}^{-1} \exp\left(\varphi n^{-2} Q(x) q'\right) \\
&= \sum_{x \in D} Q(\mathcal{T}(x)) Z_{q'}^{-1} \exp\left(\varphi n^{-2} Q(\mathcal{T}(x)) q'\right) \\
&= -\sum_{x \in D} Q(x) Z_{q'}^{-1} \exp\left(-\varphi n^{-2} Q(x) q'\right) \\
&= -e_{-\varphi}(Q_t \mid Q_{t-1} = q').
\end{aligned}$$

Thus we have that $V_\varphi(Q_t \mid Q_{t-1} = q')$ is even since $n^{-2} e_\varphi(Q_t^2 \mid Q_{t-1} = q')$ is the derivative of $(1/q') e_\varphi(Q_t \mid Q_{t-1} = q')$ with respect to φ. □

Proposition A.2. *For some $\varepsilon \in (0, \varphi\kappa)$ where $\kappa = \lim_{n\to\infty} q'/n^2$,*

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n^2} e(Q_t \mid Q_{t-1} = q') &= 2\varphi\kappa + \mathcal{O}(\varepsilon^3), \\
\lim_{n\to\infty} \frac{1}{n^2} V(Q_t \mid Q_{t-1} = q') &= 2 + \mathcal{O}(\varepsilon^2).
\end{aligned}$$

Proof. This is an immediate consequence of [11] since by trading φ in the Ising model for $\varphi n^{-2} q'$ we have the result claimed. □

Theorem A.4. *Approximately for large lattices*

$$Q_t \overset{\pi}{\in} N(0, n^2/(0.5 - 2\varphi^2)).$$

Proof. First let us see that $\pi_Q(q) = C \exp(-q^2(1 - 4\varphi^2)/(4n^2))$ satisfies the equilibrium equation. From Proposition 3.2 and from (3.1) in Section 3 we have that approximately when n is large

$$\begin{aligned}
\pi_Q(q') p_Q(q \mid q') &= C \exp\left\{-\frac{(q')^2(1-4\varphi^2)}{4n^2}\right\} \frac{1}{\sqrt{4\pi}\, n} \exp\left\{-\frac{(q - 2\varphi q')^2}{4n^2}\right\} \\
&= \frac{C}{\sqrt{4\pi}\, n} \exp\left\{-\frac{(q')^2}{4n^2} + \frac{(q')^2\varphi}{n^2} - \frac{q^2}{4n^2} + \frac{\varphi q q'}{n^2} - \frac{\varphi^2 (q')^2}{n^2}\right\} \\
&= \frac{C}{\sqrt{4\pi}\, n} \exp\left\{-\frac{q^2}{4n^2} + \frac{q^2\varphi}{n^2} - \frac{(q')^2}{4n^2} + \frac{\varphi q q'}{n^2} - \frac{\varphi^2 q^2}{n^2}\right\} \\
&= C \exp\left\{-\frac{q^2(1-4\varphi^2)}{4n^2}\right\} \frac{1}{\sqrt{4\pi}\, n} \exp\left\{-\frac{(q' - 2\varphi q)^2}{4n^2}\right\} \\
&= \pi_Q(q) p_Q(q' \mid q).
\end{aligned}$$

Evidently $C = \sqrt{(1-4\varphi^2)/(4\pi n^2)}$ and π_Q is the normal density function with $\mu = 0$ and $\sigma^2 = n^2/(0.5 - 2\varphi^2)$. $\square$

Acknowledgements

I would very much like to thank Professor Marianne Frisén and Associate Professor Aila Särkkä for good advice in my earlier work. Many thanks also to Georgi Tchilikov for fruitful discussions.

References

[1] J. Besag (1974) Spatial interaction and the statistical analysis of lattice systems. *J. Royal Stat. Soc. Ser. B Stat. Methodol.* **36** (2), 192–225.

[2] J. Besag (1991) Bayesian image restoration with two applications in spatial statistics. *Ann. Inst. Statist. Math.* **43** (1), 44–45, 50–52.

[3] J. Besag and J. Tantrum (2003) Likelihood analysis of binary data in space and time. In: *Highly Structured Stochastic Systems*, P.J. Green, N. Lid-Hjort, S. Richardson (eds.), Oxford University Press, 289–295.

[4] P. Clifford (1990) Markov random fields in statistics. In: *Disorder in Physical Systems: A Volume in Honour of John M. Hammersley*, G.R. Grimmett, D.J.A. Welsh (eds.), Clarendon Press, 19–32.

[5] M. Frisén (2003) Statistical surveillance. Optimality and methods. *Int. Statist. Rev.* **71** (2), 403–434.

[6] M. Frisén and P. Wessman (1999) Evaluations of likelihood ratio methods for surveillance: differences and robustness. *Comm. Statist. Simulation Comput.* **28** (3), 597–622.

[7] X. Guyon (1995) *Random Fields on a Network*. Springer-Verlag, New York.

[8] F. Huang and Y. Ogata (2002) Generalized pseudo-likelihood estimates for Markov random fields on a lattice. *Ann. Inst. Statist. Math.* **54** (1), 1–18.

[9] O. Häggström and K. Nelander (1998) Exact sampling from anti-monotone systems. *Statist. Neerlandica* **52** (3), 360–380.

[10] V. Ibáñez and A. Simó (2004) Empirical Bayesian estimation of the defect status on visual field longitudinal data. Submitted for publication.

[11] E. Järpe (1999) Surveillance of the interaction parameter in the Ising model. *Comm. Statist. Theory Methods* **28** (12), 3009–3027.

[12] E. Järpe (2000) Surveillance, environmental. In: *Encyclopedia of Environmetrics*, A.H. El-Shaarawi, W.W. Piegorsch (eds.), Wiley, 2150–2153.

[13] E. Järpe and P. Wessman (2000) Some power aspects of methods for detecting different shifts in the mean. *Comm. Statist. Simulation Comput.* **29** (2), 633–646.

[14] R. Kindermann and J.L. Snell (1980) *Markov Random Fields and Their Applications*. American Mathematical Society.

[15] Y. Ogata and M. Tanemura (1984) Likelihood analysis of spatial point patterns. *J. Royal Stat. Soc. Ser. B Stat. Methodol.* **46**, 496–518.

[16] L. Onsager (1944) Crystal statistics I. A two-dimensional model with an order-disorder transition. *Phys. Rev.* **65** (3–4), 117–149.

[17] E.S. Page (1954) Continuous inspection schemes. *Biometrika* **41**, 100–115.

[18] D.K. Pickard (1976) Asymptotic inference for an Ising lattice. *J. Appl. Probab.* **13**, 486–497.

[19] J.G. Propp and D.B. Wilson (1996) Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures Algorithms* **9** (1–2), 223–252.

[20] S.W. Roberts (1959) Control chart tests based on geometric moving averages. *Technometrics* **42** (1), 239–250.

[21] A.N. Shiryaev (1963) On optimum methods in quickest detection problems. *Theory Probab. and Appl.* **8** (1), 22–46.

[22] M.S. Srivastava and Y. Wu (1993), Comparison of some EWMA, CUSUM and Shiryaev – Roberts procedures for detecting a shift in the mean. *Ann. Statist.* **21** (2), 645–670.

[23] A. Särkkä (1993) *Pseudo-likelihood Approach for Pair Potential Estimation of Gibbs Processes.* Jyväskylä Studies in Computer Science, Economics and Statistics **22**, Jyväskylä.